

The European Union's IPA 2013 Programme

“Support to the Improvement of Statistical Information System” - Albania

Europe Aid/136334/IH/SER/AL

Service contract no. (CRIS) AL/IPA2013/03

Output 1a2.2:

Consultations with stakeholders

Florian Nika, Eneida Topulli and Ogerta Elezaj

Compiled by: Dragan Vukmirovic


Republika e Shqipërisë
Instituti i Statistikave

This project is funded by
The European Union

A project implemented by



 

April 2017

The Contents of this report are the sole responsibility of Ascenta and IN2 Consultants and can in no way be taken to reflect views of the European Union

Contents

PROJECT REFERENCES	3
REPORT OVERVIEW	4
COMMON CONCLUSIONS	5
RECOMMENDATIONS	9
DESCRIPTION OF REPORT OBJECTIVES	13
BACKGROUND	14
METHODOLOGICAL APPROACH	15
THE OVERVIEW	16
ANNEX 1. ABBREVIATIONS AND ACRONYMS	18
ANNEX 2: MISSION REPORT – FLORIAN NIKA	20
ANNEX 3: MISSION REPORT – ENEIDA TOPULLI	26
ANNEX 4: MISSION REPORT – OGERTA ELEZAJ	36



PROJECT REFERENCES

Project number	reference	Europe Aid/136334/IH/SER/AL
Project title	Support to the Improvement of Statistical Information System	
Contract number	AL/IPA2013/03	
Location	Albania	
Contractor	The project is implemented by a consortium led by ASCENTA IT SERVICES SRL in cooperation with IN2 Zagreb	
Contracting Authority	Central Finance and Contracting Unit (CFCU) within the Ministry of Finance	
Beneficiary	The national statistical institution of Albania (INSTAT)	
Project start date	06 th of June 2016	
Project end date	05 th of June 2018	
Project duration	24 months	

Status of document:

Submitted on	18 th of April 2017	
Version	1.0	
Status	Draft	
	For discussion	
	Approved	x

Change history:

Version	Date	Summary of Change
V1.0	18 th of April 2017	Initial version for the Beneficiary's for discussion



REPORT OVERVIEW

Project:	Support to the Improvement of Statistical Information System
Component:	Component 1 - Redesign and improve corporate statistical business processes and enhance institutional capacities in the production and usage of Agricultural and Labour Force statistics
Sub-component:	1a. Redesign and improve corporate statistical business processes
Activity:	1a.2: Review and improve core statistical business process environments
Task:	Task 2: Consultation with stakeholders
Sub-activities:	1a2.6: Separate and integrated consultations with the directorates in the central office 1a2.7: Consultation with the Ministry of Agriculture, Rural Development and Water Administration (MARDWA) 1a2.8: Consultation with the Ministry of Social Welfare and Youth 1a2.9: Consultation with the Ministry of Finance 1a2.10: Consultation with stakeholders in other similar active programs and projects 1a2.11: Report preparation
Output:	Output 1a2.2: Report on consultations with stakeholders
Experts:	Florian Nika, Eneida Topulli and Ogerta Elezaj
Date of report:	18 th of April 2017
Venue:	Tirana, Albania



COMMON CONCLUSIONS

Subcomponent 1a: Redesign and improve corporate statistical business processes

INSTAT Data

A standard model or system to save statistical data, row data (micro data) and aggregated data (macro data), does not exist. Every survey has at least two file versions of micro data (a row version and a cleaned weighted version) and a final version of aggregated data (macro data). Statistical data are saved in different formats for different surveys.

Micro data

Data collecting form also defines the format of the persisted micro data. Currently in INSTAT there are four main forms for collecting data:

- 1) **Paper Questionnaire** (Face to Face interviews)
Fulfilled paper questionnaires are then digitalized using a centralized data entry process. For data entry mainly a CSPro¹ developed data entry application is used, but, there are also some surveys that uses a Visual Studio .Net developed data entry application. The free version of CSPro saves the data entered in local text files that are later joined and exported in some statistical supported formats. In a Visual Studio developed application the model of saving entered data is not limited, but generally a centralized SQL Server database is used.
- 2) **Computer assisted personal interviewing** (CAPI) - using windows 7 laptops or Android tablets with CSPro.
CAPI exclusively uses CSPro developed applications for entering data.
- 3) **Scanners** (currently only one survey is using the scanning system)
The scanner system is using a proprietary SQL Server model to persist questionnaire design and row data that can be later exported in some statistical or text format.
- 4) There are also some different form of data collections that are currently being tested like Computer assisted telephone interviewing (**CATI**) and Computer assisted web interviewing (**CAWI**).

Currently there is no connection between MetaPlus² (supported by Statistics Sweden) and existing micro data. MetaPlus supports adding the path or the database where micro data are saved.

¹ <http://www.census.gov/population/international/software/cspro/>

² <https://www.h6.scb.se/metadatas/terms.aspx>



Data Quality

There is no specific IT system dealing with calculation of data quality figures. For every survey there are some SAS³ or SPSS⁴ scripts that calculates quality indicators.

Quality reports exists for the most part of the surveys at INSTAT but those reports are mostly for internal use. With implementation of Single Integrated Metadata Structure (SIMS) project in INSTAT, all quality standards and templates are going to be defined. Currently all templates and standards are defined for the Labour Force Survey (LFS), (HBS), Business Register (BR) and Structural Business Survey (SBS). Euro-SDMX Metadata Structure (ESMS) implementation is finalizing all statistical activity quality templates, ESS Standard for Quality Reports Structure (ESQRS) has not started yet.

There is no overlap between SIMS and MetaPlus as SIMS is dealing with Referential Metadata and MetaPlus is dealing with Structure Metadata. Both systems are going to be important sources for the data warehouse.

Currently there is no connection between MetaPlus and existing micro data. MetaPlus supports adding the path or the database where micro data are saved.

Taking in consideration all this different ways of data collecting and persisting that are used or going to be used in the future, a well-defined model for saving (or final exporting) micro data will be crucial. This model should support versioning with comments for every change. If the data warehouse is going to load micro data it is important to have standardized sources for this kind of data.

Subcomponent 1b: Enhance institutional capacities in the production and usage of Agricultural statistics

From consultations with stakeholders and analysis of the current situation of agricultural statistics, the summarized main issues, identified by this mission that need to be addressed in the future, are as follows:

- **Strengthening inter-Institutional cooperation**

This is a requirement of all institutions that are part of the statistical in Albania. INSTAT requires timely response and quality from statistical agencies, while the agencies need to have flexibility communicating with INSTAT for micro data and time series for policy making. Also, the necessity of strengthening the role of INSTAT as methodological leader for statistical agencies emerges.

³ https://www.sas.com/en_us/software/analytics/stat.html

⁴ <https://www.ibm.com/analytics/us/en/technology/spss/>



- **Strengthening Capacities**

Institutional capacity of statistical system in general and that of agricultural statistics in particular need quantitative and qualitative strengthening

- **Establishment of the farm register**

The Farm Register is a necessary and expected instrument for initiating other statistical activities for agriculture.

In establishing the farm register, special importance must be given to Unique Farm Identification Code, as a tool that can ensure the cooperation and update of farm register with existing and future registers (Livestock matriculation register, the register of agricultural supporting schemes, cadastre of vineyard etc. set up and administered by MARDWA or Agency for Rural Development Albania (ARDA)).

- **Increase the number and frequency of agricultural surveys**

This need comes from the shortcomings identified during the process of drafting and monitoring of agricultural policies and strategies, as well as for meeting the requirements of EUROSTAT and growing demands of other users

- **Strengthening administrative data**

Administrative data are a very important and accurate source for statistics. Every statistical system in every country tends to be oriented toward administrative data. In Albania, strengthening the role and use of administrative data is related with the establishment and improvement of aforementioned registers. Still, considering actual terms and what is defined as administrative data in section b.2 of this report (Annex 3), MARDWA needs assistance in designing a methodology for improving the quality and effectiveness of this important source of agricultural statistics in Albania. There is a need for establishment of a methodology for data collection, processing and dissemination of data that includes also instruments for forecasting and evaluation for agriculture and livestock production during and in the end of agricultural year.

- **Integrated Statistical Information System**

Agriculture and other sector strategies are generally multi sectoral. For ensuring the success on their design, implementation and monitoring there is a need for a Coordination Mechanism or establishment of an Integrated



Information System. This is an initiative undertaken by the current project and congratulated by stakeholders.

Subcomponent 1c: Enhance institutional capacities in the production and usage of Labour Force statistics

- General capacity building regarding the theoretical and practical knowledge in time series analyses with a focus on stochastic models and seasonal adjustments.
- More extensive training in SPSS/SAS (parameterized macros) using real LFS data sets for data management and time series analyses;
- General capacity building regarding the Data Management of Statistical Information System including data sources linkage, data processing and data quality.
- Active participation in all the training sessions of representatives of National Employment Service and Ministry of Social Welfare that are in charge of labour statistics.
- The representatives from Ministry and National Employment Service require actively participating in INSTAT training for increasing their capacity in data analyses and in using SPSS software for data processing.

RECOMMENDATIONS

INSTAT Data Management System

Recommendation is to increase the staff capacities via a data management training which should effectively address the growing and diverse needs for skills and competencies in INSTAT in both its content and methods for the following topics:

- Data processing and data integration
- Quality and access to data
- Metadata System
- Examples of Big Data uses in European Statistical System

Also, as a request from other institutions, the recommendation including the training sessions participants from different INSTAT units dealing with statistics production (not only LFS unit representatives) and participants from other institutions working in statistical units.

Regarding the software used during LFS data processing, there is a need to build statistical capacity for the staff to develop SPSS parameterized procedures and increase skills in writing their own macros to increase productivity and improve LFS data management. The methodology unit is using SAS procedures during weighting phase and also some procedures available by EUROSTAT for checking micro data before transmitted are written in SAS, training in SAS are important for the sector. The staff has limited knowledge in SAS. Also, further discussion are needed to be done with the IT team for the software licenses plan, because SAS/SPSS are not open source and in case of limited budget training in R (open source statistical package⁵) should be considered for the staff. At the current moment, it is not decided yet if SIDA project will provide INSTAT with the required licenses for SAS/SPSS software.

Recommendations for improving the Agriculture statistics

- After the establishment of the new statistical farm register by the agricultural census, expected after the process of calibration, a reform on agriculture statistics is necessary in terms of conducting special surveys for sensitive indicators; the sample to be objective-specific; the frequency should vary according to the objective and of course the indicators should be revised to be more compliant with EU requirements.
- Although, despite the outcome of the census calibration to move ahead statistical activities in agriculture, alternative sources must be explored and exploited such as:
 - Development of an action plan to address priority problems,
 - Improve the collection and processing of administrative data,

⁵ <https://www.r-project.org/>



- Development of metadata,
 - Analysis of the results of the annual agricultural survey conducted on the basis of population generated by the census of agriculture and calibration survey to explore the possibilities of improving and updating the database of the census and the establishment of farm register
 - Review and improve the methodologies on existing data collection
 - Design and piloting of the establishment of Farm Accountancy Data Network (FADN), etc.
- The focus of agricultural statistics in Albania is put on quantitative data on agricultural activity during the year. It is important to plan the production of qualitative data and statistics on socio-economic development of rural areas that can provide a deeper understanding of the impact of Agricultural and rural development policies.
 - Strengthen and increase capacity in the system of agricultural statistics is a necessity. SIDA Project and the current IPA project can unite their energies to formulate a specific long-term training plan for agricultural statisticians to ensure the continuous training of experts in this field.
 - Building a strategy for the publication of statistics, to address the necessary indicators to be published, their frequency, coherence, necessary explanations for users etc., is an important tool for improving the dissemination process.
 - It is also necessary to draft a plan of action for the evaluation of any stage to the statistical business process and quality control.
 - Improvements of statistical business processes should be guided by the framework of an Information System that needs to be drafted soon, and implemented after the establishment of all its components.



Recommendations for improving the LFS

LFS time series analyses:

Recommendation is to increase the LFS staff capacities via a time series training which should effectively address the growing and diverse needs for skills and competencies in INSTAT in both its content and methods for the following topics:

- Characteristics of time series data
- Stochastic models
- Forecasting of several time series
- Examples and applications to the LFS survey (SPSS examples)

Other recommendations

Metadata

The metadata system (MetaPlus) will be one of the most important sources for the data warehouse system, thus, recommendation is to start defining and developing metadata retrieval procedures from MetaPlus.

Quality

Data warehouse has to be primary source for quality indicator reports. Also when designing the data warehouse model consider adding some quality variables that can be calculated during ETL⁶ process of the data warehouse.

Micro data

Taking in consideration different ways of data collecting and persisting that are used or going to be used in the future, a well-defined model for saving (or final exporting) micro data will be crucial. This model should support versioning with comments for every change. If the data warehouse is going to load micro data it is important to have standardized sources for this kind of data.

Macro data

Currently macro data (aggregated data) are saved in different formats like: excel files, SPSS files, SAS files and sometimes all three formats exists for the same data.

There are two main aggregated statistical data models that are generally used today for dissemination and data exchange:

- SDMX
SDMX is the Eurostat data model for exchanging statistical data, with a lot of tools developed from Eurostat to export and transfer data in this model.

⁶ <http://datawarehouse4u.info/ETL-process.html>

- **Nordic Data Model**

This is the model PxWeb⁷ is using for disseminating tabled and chart data

Both model are interchangeable and can be used to save aggregated data for different proposes. INSTAT IT staff is doing, or will do in the future, huge redundant work to export aggregated data in both models: SDMX (for exchanging statistical data and metadata with Eurostat) and NDM (for web data dissemination using PxWeb).

In order to reduce the effort for data transformation recommendation is to adding SDMX and NDM format support to the data warehouse reporting system.

Data Center

As SQL Server is decided to be used as primary DBMS for the data warehouse, selecting the right version is important and AKSHI can support getting a standard edition license.

Recommendation: Coordination between related tasks from both projects is very important, especially for tasks that have dependencies. Having a data center in place, a metadata system implemented and all needed software licenses is crucial for the IT components of the IPA 2013 project.

MetaPlus integrated with the website

SIDA has plans to update the CMS and the design of INSTAT website. INSTAT is using Umbraco Content Management System (CMS) for its website. This CMS has been chosen because it can be extended and integrated with own developed C# web applications.

This output will involve some integration development between Umbraco and MetaPlus. Having information on this integration could be a good start to define a way of communication between MetaPlus and the data warehouse.

Structured and standardized data storage system for statistical data

Structured and standardized data storage system for statistical data is another output of SIDA Phase 4 project that is with interest. As suggested above, having a standardized data storage system for statistical data is a very important.

This output is not planned yet and recommendation is to coordinate plans for this output and the related IPA 2013 outputs.

SIDA representatives concerns

Variables and value sets standardization is one of the most difficult task when designing a good metadata system for statistics. Still a huge work needs to be done

⁷ http://www.scb.se/sv_/PC-Axis/Programs/PX-Web/



in INSTAT in order to standardize every variable used in surveys. This is a problem that the project of data warehouse implementation will face sooner or later.

Working together (SIDA, IPA 2013 and INSTAT experts) is very important to successfully finalize this task.

DESCRIPTION OF REPORT OBJECTIVES

Objectives:

Expected results related to Component 1, subcomponent 1a, is related with upgrading statistical business process architecture with improved data management, meeting European standards and practices, designed and implemented in INSTAT.

Expected results related with the Task 2: Consultations with stakeholders formalized in this report is to assist the implementation of business process model expected results related with the Activity 1a.2: Review and improve core statistical business process environments, Task 2: Consultation with stakeholders is presented as this Output 1a.2.2: Report on consultations. This report is designed based on an inventory of the actual processes, taking into account the expectations with respect to future changes (Output 1a.2.1: *Review of the processes & data flows and necessary improvements in INSTAT*) and follows Output 1a.1 Blueprint of a corporate statistical business process architecture aligned with the Generic Statistical Business Process Model (GSBPM)⁸.

The **main objectives** of this report are:

- Review of definitions, data flows and protocols between INSTAT and main stakeholders
- Recommendations for the future improvements on data flows and protocols.

The **specific objectives**:

- Report on consultations with emphasis on the technical specifications for statistical IT applications related with metadata and quality:
 - Consultation with the stakeholder in other similar active programs and projects, especially with SIDA representatives
 - Analyze and recommendations for future development metadata system in INSTAT: MetaPlus or Eurostat solution
- Report on consultations with emphasis on the peculiarity of the agriculture statistics:
 - Consultation with the stakeholders: MARDWA, Ministry of Finance and other similar active programs and projects: SIDA project, other IPA projects and other donors

⁸ <http://www1.unece.org/stat/platform/display/GSBPM/GSBPM+v5.0>



- Report on consultations with the stakeholders and assessment of future needs of MARDWA and Ministry of Finance related with Agriculture statistics.
- Report on consultations with emphasis on the specificity of the Labour market:
 - Consultation with the stakeholders: Ministry of Social Welfare and Youth, MARDWA, Ministry of Finance and other similar active programs and projects: SIDA project, other IPA projects. Basic document for consultation is Project output 1a2.1. Report: Review of the processes & data flows and necessary improvements in INSTAT.
 - Consultation with LFS team in INSTAT and other relevant data providers about training needs in the fields of LFS.

This report is the basis for the following activities to be undertaken under the Component 1, but also the basis for the IT systems developments as identified under Component 2 and some activities under Component 3.

BACKGROUND

The report is related to the work within the IPA 2013 Project: Support to the Improvement of Statistical Information System (Project). Project is conceptualized into three main components with large number of activities, sub-activities and tasks having a different degree of complexity and largely depending on a number of various factors, of internal and external nature.

Component 1 - Redesign and improve corporate statistical business processes and enhance institutional capacities in the production and usage of Agricultural and Labour Force statistics

Component 2 - Develop and test system for upgraded corporate ICT infrastructure with inclusion of establishment and pilot implementation of centralised Data Warehouse(s)

Component 3 - Strengthen communication function and statistical coordination role of INSTAT with focus on dissemination and quality management.

The purpose of this Project is to enhance the institutional capacities in the Albanian Statistical System (ASS), and in particular in INSTAT, for streamlining the statistical processes and development of a redesigned IT architecture based on the data warehouse concept, enabling efficient and effective production, analysis and dissemination of official statistics meeting the requirements of the acquis in statistics and following best practices as implemented in the European Statistical System (ESS).



This report on consultations with stakeholders is an output (Output1a2.2) for the:

- **Component 1:** Redesign and improve corporate statistical business processes and enhance institutional capacities in the production and usage of Agricultural and Labour Force statistics;
 - **Sub-component 1a:** Redesign and improve corporate statistical business processes;
 - **Activity 1a.2:** Review and improve core statistical business process environments
 - **Task 2:** Consultation with stakeholders.

METHODOLOGICAL APPROACH

The methodological approach for creating this Report was based on five sub-activities:

- 1a2.7: Consultation with the Ministry of Agriculture, Rural Development and Water Administration (MARDWA)
- 1a2.8: Consultation with the Ministry of Social Welfare and Youth
- 1a2.9: Consultation with the Ministry of Finance
- 1a2.10: Consultation with stakeholders in other similar active programs and projects
- 1a2.11: Report preparation

The **first step** was to review the input documents and other necessary information:

- Law no. 9180, dated 05.02.2004 "official statistics",
- Official Statistics Programme for the period 2017-2021 (OSP 2017-2021)
- IPA 2014 Multi-beneficiary statistical cooperation programme, MISSION REPORT - DRAFT, Pilot Project 18, Developing a Metadata System, May, 2016,
- Project outputs/reports:
 - Blueprint of a corporate statistical business process architecture aligned with the Generic Statistical Business Process Model (GSBPM)
 - Assessment of the current situation in agricultural statistics, SFR quality and methodology and TA programs,
 - Assessment of the inter-institutional communication and reporting protocols in the area of Agricultural statistics
 - Review of the processes & data flows and necessary improvements in INSTAT
 - Improvements on data collection methodologies, standardisation and modern tools
- Report on sector review of the Business statistics in Albania, 2015



- Sector review of Implementation of the Labour Force Survey in Albania, Draft report ver. March, 2016,
- Report on Limited Peer Review of the Institute of Statistics of Albania, 2015
- Short Term Statistics – STS 2014, Technical Report, INSTAT, www.instat.gov.al
- Quality reports and definitions,
- Non-key expert (NKE) mission reports

Second step: Consultation with INSTAT staff before interviews with stakeholders.

Third step: Consultation with external stakeholders.

Fourth step: Consultation with INSTAT staff – after interviews with stakeholders for the clarification

Five step: Analysis of consultations.

Final step: Report preparation - define recommendations for the improvements definitions, data flows and protocols in INSTAT. Report is prepared by the Team leader/Key expert 1 on the basis of the mission reports of NKEs: Florian Nika, Eneida Topulli and Ogerta Elezaj.

THE OVERVIEW

The main objective of the Project is to support INSTAT to update and upgrade in a structured way the processes documentations in the framework of enhanced Quality management. The targeted result is the improvement of the efficiency and effectiveness of the INSTAT's operation and at the same time preparation the Institute for the establishment of centralised data warehouse and for the introduction of the new information technologies and methodologies. The statistical business process environments in INSTAT are reviewed. Clearly defined business process is the essential for the successful production of the official statistics. The development of the business process is the base for the implementation of this Project and for the further modernization of the statistical production processes in INSTAT.

For the full implementation of the business process model, three main activities are undertaken:

- The review of the existing processes in INSTAT *with focus on data flows*,
- The review of data collection methodologies and registers, and
- The identification of necessary improvements.

The view of the core statistical processes is envisaged through this project. Executed tasks are the review of the processes *with focus on data flows*, identification of necessary improvements, and assistance with their implementation. Further on, the project can help INSTAT to update and upgrade in a structured way the processes'



documentations in the framework of enhanced Quality management. The targeted result is the improvement of the efficiency and effectiveness of the INSTAT's operation and at the same time preparation the Institute for the establishment of centralised data warehouse(s) and for the introduction of the new information technologies and methodologies. Also the statistical business process environments shall be reviewed and upgraded.

The frame for the basic Inventory of the existing processes and data collection methodologies in INSTAT is Official Statistics Programme for the period 2017-2021 (OSP 2017-2021) and the Statistical requirements compendium (Compendium)⁹. Five OSP 2017-2021 is the basic document that provides the production of statistical data by the National statistical system needed for the observation of economic, social and environmental situation in the Republic of Albania, respecting the guarantee of the implementation of statistical principles provided in law and in the European Statistics code of Practice.

INSTAT's OSP 2017-2021 follows the Compendium. The structure of the 2016 edition of the Compendium generally follows the structure of the European Statistical Program 2013-2017 and the corresponding Annual Work Programs of Eurostat. For candidate countries such as Republic of Albania, the compendium is used to assess compliance level with the EU acquis in the area of statistics. The presentation of the priority areas and modules is made from the perspective of a statistical data provider in a country. **It means that in general only modules for which there is a legal or methodological basis that influences the way statistics are produced in countries being part of the European Statistical System** or for which there are obligations to provide data to Eurostat are included.

⁹ <http://ec.europa.eu/eurostat/documents/3859598/7535696/KS-GQ-16-003-EN-N.pdf/d3f54ab4-97af-4f53-be59-45aca3ecfb80>



ANNEX 1. ABBREVIATIONS AND ACRONYMS

AKSHI - Agjencia Kombëtare e Shoqërisë së Informacionit / National Agency for Information Society

ARDA - Agency for Rural Development Albania

ASS - Albanian Statistical System

BR – Business Register

CAPI – Computer assisted personal interviewing

CATI – Computer assisted telephone interviewing

CAWI – Computer assisted web interviewing

CFCU - Central Finance and Contracting Unit

CMS - Content Management System

Compendium - The Statistical requirements compendium

DBMS - Database Management Systems

ESMS - Euro-SDMX Metadata Structure

ESQRS - ESS Standard for Quality Reports Structure

ESS - European Statistical System

EU – European Union

EU SILC - European Union Statistics on Income and Living Conditions

FADN - Farm Accountancy Data Network

GSBPM - Generic Statistical Business Process Model

HBS – Household Budget Survey

INSTAT - Institute of Statistics of Albania

ICT – Information Communication Technology

IT – Information Technology

IPA - Instrument for Pre-accession Assistance

IPA 2014 - Multi-beneficiary statistical cooperation programme



LFS - The Labour Force Survey

LSMS – Living Standards Measurement Survey

MARDWA - Ministry of Agriculture, Rural Development and Water Administration

NDM - Network Data Mover

NES - National Employment Service

NKE - Non-key expert

OSP 2017-2021 - Official Statistics Programme for the period 2017-2021

SBS - Structural Business Survey

SDMX - Statistical Data and Metadata eXchange

SFR – Statistical Farm Register

SIDA – Swedish International Development Cooperation Agency

SIMS – Single Integrated Metadata Structure

STS - Short Term Statistics

TA – Technical Assistants



ANNEX 2: MISSION REPORT – FLORIAN NIKA

Project Title Contract no.	Support to the Improvement of Statistical Information System – Albania EuropeAid/136334/IH/SER/AL
Beneficiary	Statistical Institute of Albania (INSTAT)
Name of Short-Term Expert	Florian Nika
Component	Redesign and improve corporate statistical business processes; and enhance institutional capacities in the production and usage of Agricultural and Labour Force statistics
Sub-Activity	1a2.10 Consultation with stakeholders in other similar active programs and projects
Date of mission	17-27 of January 2017
Date of report	February 02, 2017

1. Description of mission objective

- Expertise advice and technical assistance for preparation Output 1a2.2: Report on consultations with emphasis on the technical specifications for statistical IT applications related with metadata and quality
- Consultation with the stakeholder in other similar active programs and projects, especially with SIDA representatives.
- Report on consultations with the stakeholders – analyze and recommendations for future development metadata system in INSTAT: MetaPlus or Eurostat solution

2. List of meetings held during the mission

During the mission the following meetings with Beneficiary were held:

- Meeting with IT directorate

Participant	Email	Role
Helda Mitre	hmitre@instat.gov.al	IT director
Alma Kondi	akondi@instat.gov.al	Head of Quality Statistics Sector
Olta Kodra	okodra@instat.gov.al	Head of Database and Metadata Sector



- Meeting with SIDA representative

Participant	Email	Role
Birgitta Mannfelt	Birgitta.mannfelt@scb.se	Senior Advisor

- Meeting with SIDA representative

Participant	Email	Role
Marjana Gorica	mgorica@instat.gov.al	Project Coordinator

3. Results achieved during the mission

- Report on consultation with stakeholder.
- General overview of current metadata system (MetaPlus) in INSTAT
- General overview of current data quality system in INSTAT
- General overview of current statistical data systems in INSTAT

Metadata

SIDA project has invested on building a metadata system in INSTAT using MetaPlus, a system already tested and used at statistics Sweden.

MetaPlus

Administrative sources are documented and inputted in the MetaPlus system. INSTAT has started documenting administrative sources because it is simpler to standardize variables and value sets for this kind of data.

A good job is done so far to define metadata for some of INSTAT surveys, registers and censuses like: STS, LFS, Business Register and Census 2011 but none of the statistical activity of INSTAT is persisted in MetaPlus yet.

SIDA has already planned a training session that will take place on February, to help INSTAT defining metadata for all its surveys.

MetaPlus is developed in C#.Net¹⁰ and the work to adopt the code to INSTAT needs is already finished. SQL Server is used for data persistence in MetaPlus, and currently, there is no API to access or retrieve metadata information. The client application of MetaPlus is a desktop application that needs to be installed in the local computers of the users that are going to use the system. INSTAT, in order to provide

¹⁰ <https://msdn.microsoft.com/en-us/library/z1zx9t92.aspx>

understandable information about its statistics, has plans to put metadata information directly into the website. It is not known how it is going to take place the communication between the website and MetaPlus. My opinion is that it is going to be some metadata retrieval directly from MetaPlus SQL Server database. Having a good knowledge on MetaPlus data model is crucial to communicate with MetaPlus.

The metadata system (MetaPlus), in my opinion, will be one of the most important sources for the data warehouse system, thus, my recommendation is to start defining and developing metadata retrieval procedures from MetaPlus.

Data Quality

There is no specific IT system dealing with calculation of data quality figures. For every survey there are some SAS or SPSS scripts that calculates quality indicators.

Quality reports exist for the most part of the surveys at INSTAT but those reports are mostly for internal use. With implementation of SIMS project in INSTAT, all quality standards and templates are going to be defined. Currently all templates and standards are defined for LFS, HBS, BR and SBS). ESMS implementation is finalizing all statistical activity quality templates, ESQRS has not started yet.

There is no overlap between SIMS and MetaPlus as SIMS is dealing with Referential Metadata and MetaPlus is dealing with Structure Metadata. Both systems are going to be important sources for the data warehouse.

My recommendation is to use data warehouse as primary source for quality indicator reports. Also when designing the data warehouse model consider adding some quality variables that can be calculated during ETL process of the data warehouse.

INSTAT DATA

A standard model or system to save statistical data, raw data (micro data) and aggregated data (macro data), does not exist. Every survey has at least two file versions of micro data (a raw version and a cleaned weighted version) and a final version of aggregated data (macro data). Statistical data are saved in different formats for different surveys.

Micro data

Data collecting form also defines the format of the persisted micro data. Currently in INSTAT there are four main forms for collecting data:

- **Paper Questionnaire** (Face to Face interviews).
 - Fulfilled paper questionnaires are then digitalized using a centralized data entry process. For data entry mainly a CSPro developed data entry application is used, but, there are also some surveys that use a Visual Studio .Net developed data entry application. The free version of CSPro saves the data entered in local text files that are later joined and

exported in some statistical supported formats. In a Visual Studio developed application the model of saving entered data is not limited, but generally a centralized SQL Server database is used.

- **CAPI** (using windows 7 laptops or Android tablets with CSPro).
 - CAPI exclusively uses CSPro developed applications for entering data.
- **Scanners** (currently only one survey is using the scanning system).
 - The scanner system is using a proprietary SQL Server model to persist questionnaire design and row data that can be later exported in some statistical or text format.
- There are also some different form of data collections that are currently being tested like **CATI and CAWI**.

Currently there is no connection between MetaPlus and existing micro data. MetaPlus supports adding the path or the database where micro data are saved.

Taking in consideration all this different ways of data collecting and persisting that are used or going to be used in the future, a well-defined model for saving (or final exporting) micro data will be crucial. In my opinion, this model should support versioning with comments for every change. If the data warehouse is going to load micro data it is important to have standardized sources for this kind of data.

Macro data

Currently macro data (aggregated data) are saved in different formats like: excel files, SPSS files, SAS files and sometimes all three formats exists for the same data.

There are two main aggregated statistical data models that are generally used today for dissemination and data exchange:

- **SDMX**
SDMX is the Eurostat data model for exchanging statistical data, with a lot of tools developed from Eurostat to export and transfer data in this model.
- **Nordic Data Model**
This is the model PxWeb is using for disseminating tabled and chart data

Both model are interchangeable and can be used to save aggregated data for different proposes. INSTAT IT staff is doing, or will do in the future, huge redundant work to export aggregated data in both models: SDMX (for exchanging statistical data and metadata with Eurostat) and NDM (for web data dissemination using PxWeb).

In order to reduce the effort for data transformation I recommend adding SDMX and NDM format support to the data warehouse reporting system.

Consultation with SIDA representatives



SIDA Phase 4 Project Outputs

Data Quality, evaluation and imputation

Data quality, evaluation and imputation is one of the output listed in the area of General Agricultural Statistics from SIDA Phase 4 project. Currently nothing is done to accomplish this output and SIDA representative does not think there will be any conflicts between this output and related IPA 2013 output.

Production of statistics in a reliable IT system improving timeliness and process quality

Production of statistics in a reliable IT system improving timeliness and process quality is another output listed in the area of Information Communication Technology from SIDA Phase 4 project that it is with interest for the IPA 2013 project. This area of investment from SIDA is related only to hardware and networking equipment and systems. There is no specific system related to statistics production. It is decided that all hardware and network procurement will be related to the IPA 2013 project needs for hardware and network configuration, especially the Data Centre specifications.

Data Center

Currently the procurement procedures for buying all needed equipment for the data center has not been started yet. SIDA representatives are still waiting for the server specifications from INSTAT. Some consultation with National Agency for Information Society (AKSHI) has taken place to confirm the licenses that can be obtained from them, also there are some legal procedures that oblige every IT related procurement to get a confirmation from AKSHI.

As SQL Server is decided to be used as primary Data Warehouse Database Management Systems (DBMS) for the data warehouse, selecting the right version is important and AKSHI can support getting a standard edition license.

Recommendation: Coordination between related tasks from both projects is very important, especially for tasks that have dependencies. Having a data center in place, a metadata system implemented and all needed software licenses is crucial for the IT components of the IPA 2013 project.

MetaPlus integrated with the website

SIDA has plans to update the CMS and the design of INSTAT website. INSTAT is using Umbraco CMS for its website. This CMS has been chosen because it can be extended and integrated with own developed C# web applications.

In my opinion this output will involve some integration development between Umbraco and MetaPlus. Having information on this integration could be a good start to define a way of communication between MetaPlus and the data warehouse.



Structured and standardized data storage system for statistical data

Structured and standardized data storage system for statistical data is another output of SIDA Phase 4 project that is with interest. As suggested above, having a standardized data storage system for statistical data is a very important.

This output is not planned yet and my recommendation is to coordinate plans for this output and the related IPA 2013 outputs.

SIDA REPRESENTATIVES CONCERNS

Variables and value sets standardization is one of the most difficult task when designing a good metadata system for statistics. Still a huge work needs to be done in INSTAT in order to standardize every variable used in surveys. This is a problem that the project of data warehouse implementation will face sooner or later.

Working together (SIDA, IPA 2013 and INSTAT experts) is very important to successfully finalize this task.

4. Problems encountered during the mission

No problems have been encountered during the mission.

5. Recommendations for improving the sector

TO BE ANALYSIS

- Recommendations on future development of MetaPlus integration with the data warehouse
- Recommendations on future development of statistical data system integration with the data warehouse
- Recommendations on future development of data quality system integration with the data warehouse

6. Plans for future missions

Future missions will be planned as stated in Inception report

7. Attachments (reports, documents, presentations, etc.)

Report on consultation document



ANNEX 3: MISSION REPORT – ENEIDA TOPULLI

Project Title Contract no.	Support to the Improvement of Statistical Information System – Albania EuropeAid/136334/IH/SER/AL
Beneficiary	Statistical Institute of Albania (INSTAT)
Project implemented by	The project is implemented by a consortium led by ASCENTA IT SERVICES SRL in cooperation with IN2 Zagreb
Name of Short-Term Expert	Eneida Topulli
Component	Component 1 - Redesign and improve corporate statistical business processes and enhance institutional capacities in the production and usage of Agricultural and Labour Force statistics
Sub Component	Sub-component 1a: Redesign and improve corporate statistical business processes
Activity	Activity 1a.2: Review and improve core statistical business process environments
Task	Task 2. Consultation with stakeholders
Date of mission	15 February -7 March
Date of report	14 March

1. Description of mission objective

- Expertise advice and technical assistance for preparation Output1a2.2: Report on consultations with emphasis on the peculiarity of the agriculture statistics.
- Consultation with stakeholders: MARDWA, Ministry of Finance and other similar active programs and projects: SIDA project, other IPA projects and other donor programs like GIZ. Basic document for consultation is Project output 1a2.1. Report: Review of the processes & data flows and necessary improvements in INSTAT and future stakeholders needs.
- Report on consultations with the stakeholders and assessment of future needs related with Agriculture statistics.

2. List of meetings held during the mission



During the mission the following meetings were held:

Participant	Email	Role
Mr. Alban Cela	acela@instat.gov.al	Director of Directorate of Agricultural and Environmental Statistics (INSTAT)
Mr. Flamur Hysi	Flamur.Hysi@bujqesia.gov.al	General Director of Agricultural Policies, (MARDWA)
Mrs. Qanie Fetahu	Qanie.Fetahu@bujqesia.gov.al	Specialist of Agricultural Statistics, MARDWA
Mrs. Birgitta Mannfelt	birgitta.mannfelt@scb.se	Senior advisor of SIDA project-Partnership in statistics in Albania, Phase IV
Ms. Alina Jahn	AJahn@instat.gov.al	Expert of Agricultural and Environmental Statistics of CIM
Ms. Elida Fara	Elide.fara@financa.gov.al	Head of Unit Macroeconomic analyses and statistics at Ministry of Finance

3. Results achieved during the mission

3.1. Overview of statistical processes from stakeholder's perspective

This mission under component 1 is focused on the identification of issues and needs of relevant institutions in the production and use of statistics. The scope is to efficiently focus project resources in achieving the main objective of this component that is enhancing institutional capacities in the production and use of agricultural statistics.

For this purpose, a series of meetings were held with relevant persons who are producers and users of agricultural statistics, as well as representatives of projects that have assisted and have on their agenda similar activities.

Discussions were focused on issues and future needs on data flow and statistical business.

A. Inter-Institutional Cooperation

INSTAT is the authorized institution in coordinating the entire statistical system in Albania. Its relationship with other statistical agencies and in particular with MARDWA, is defined by law and other bylaws.

A Memorandum of Understanding is also signed between the two institutions to facilitate the process of frequent information exchange.



Both institutions agree that their tasks and responsibilities are clear and well defined. However Inter-institutional communication is still not fully efficient. According to INSTAT and MARDWA, they have clear agenda but they face difficulties in implementation.

INSTAT needs that statistical agencies improve data quality and response time. On the other hand, MARDWA needs from INSTAT to strengthen its role as methodological leader in other processes that are not defined in formal agreements between them (case of administrative data).

Also, MARDWA and the Ministry of Finance, seen in the role of users of various indicators, state the need for a more flexible communication in order to enable the data availability by INSTAT with reduced bureaucratic procedures. This is a common problem that users face when they need more detailed information than standard format of publication available, necessary for timely analysis for decision making.

A Memorandum of Understanding is also signed to determine the responsibilities between INSTAT, Bank of Albania and the Ministry of Finance for the preparation, distribution, transmission, sharing and exchange of data.

b. Production of statistics

b.1. Primary research

INSTAT has constantly adapted the European standards in the implementation of statistical surveys in terms of survey type, frequency, indicators, methodology used, processing and analysis techniques, publication, etc.

Primary research in Agriculture has faced a transition period, conditioned by changes of organizational structure of responsible institutions, and significant discrepancies of the Agricultural Census Output with data published previously by MARDWA.

The two institutions together, are performing an annual statistical survey for 2016, the first annual survey after the last conducted by the Ministry of Agriculture in 2012. Up to 2012, MARDWA has been the statistical agency responsible for the production and publication of agricultural statistics. Several agricultural surveys were conducted with methodologies established and consolidated with the support of an USAID project and through United States Experts of National Agricultural Statistical Service and Agricultural Department.

Although methodologies and surveys have been drawn up and accepted for the actual terms, MARDWA states it has always felt the need of having more specific and frequent surveys, with necessary indicators to meet the increasingly demanding users needs, and, on the other hand, to adapt to the requirements of EUROSTAT, a need that will be returned to obligation considering Albania's aspiration to be acceded to EU. The Methodology for sample surveys, especially for annual survey of



agriculture, was the only solution for the time being, but not the best one, since one frame with the same periodicity is impossible to be fully representative for all agricultural indicators. It was designed to be fully representative mainly for agricultural land and production, orcharding and livestock numerical situation and production.

SIDA project-Partnership in statistics in Albania, Phase IV, has had in its agenda a long term assistance for agricultural surveys, planned in terms of a more favourable terrain.

For this purpose, the Census of Agriculture was expected, which would serve as basis for the realization of various and frequent surveys. Currently, INSTAT, with the support of current IPA project is working with the process of Census calibration. INSTAT has conducted a survey with more than 10,000 households to enable this process.

SIDA project has already deviated the line of activities from agricultural statistics, and is providing expertise in surveys that are linked to aquaculture and greenhouses but from the environmental prospective. This decision of SIDA project is made considering also the agenda of the current IPA 2013 project: Support to the Improvement of Statistical Information System – Albania and trying not to duplicate the activities. The latter is offering expertise for the process of agricultural census calibration and will provide support in the future for capacity building in the production and use of agricultural statistics.

Meanwhile, waiting for the preparation of a more favourable terrain for the further consolidation of agricultural statistics, the expertise of the CIM, with focus on assistance in agricultural and environmental statistics, is currently focusing on environmental statistics and on strengthening and improving data collection system for agricultural prices.

b.2 Secondary Research

The production process for administrative data in general is the responsibility of the statistical agencies. In this aspect INSTAT plays mainly the role of the user.

MARDWA administrative record consists of a yearly basic register that includes information on main agricultural indicators like mechanic tools, irrigation of land, area with crops and orcharding, numerical situation of orcharding and livestock, as well as yield and production for each culture.

The register is furnished by data provided from forecasts or estimates of extension service experts in Regional Offices of Agriculture in cooperation with the local government agronomist, small scale surveys with contact farmers, etc.

Besides the fact that these data in this form of collection are not purely administrative data, MARDWA needs to apply a standardized process for all regions in order to



minimise possible human errors during the process of forecasting and evaluation of agricultural and livestock products. Administrative data are collected and processed in a format (Excel File) that is not suitable for further processing and analysis. For this purpose MARDWA requests for assistance in establishing a methodology for improvement and consolidation of collection and processing of administrative data.

INSTAT, as user of these data, needs improvement of their quality and quantity.

Actually the collection of administrative data is an initiative of MARDWA for its own needs in the process of designing and monitoring agricultural policies and it is not included in any legal by-law that regulates the relationship or responsibilities of both institutions. Nevertheless, these data are an alternative source for agricultural statistics that have fulfilled the needs of many users, including INSTAT.

MARDWA needs methodological assistance from INSTAT in this context and INSTAT should supervise each stage of production of administrative data to meet its requirements.

c. Dissemination

INSTAT publishes statistics in two web portals: the first is the official website of INSTAT, where the data are published in Excel, PDF, and SPSS format, and the second one is the web-URL <http://databaza.instat.gov.al> that includes detailed data for all sectors of the economy. A calendar of publications is available on the website to inform users about the types of surveys, periodicity, or indicators that will be published.

Published statistics are heterogeneous and generally meet the needs of users and institutions for information. Currently agricultural statistics dissemination consists on the publication of key indicators that are related mainly with planted area, number of heads of livestock, agricultural and livestock products, areas and products in greenhouses, livestock by-products and the export-import statistics.

Indicators published for agriculture are insufficient to both institutions needs for their activities, but also for the users' requirements. MARDWA has great difficulties caused by the lack of data especially during the process of monitoring strategies or agricultural policies.

Limited published indicators in agriculture are an issue that is not related to the dissemination but to the process of production for the reasons mentioned earlier in previous sections of this report.

Results of the calibration of agricultural census and annual survey 2016 are expected to integrate agricultural statistics with the publication of all sectors.



INSTAT has recently launched in the official website the user satisfaction survey. This survey serves to test the level of satisfaction of users regarding the quality of published statistics, ease of access, frequency, etc.

Its results are expected to be an important basis for building a strategy for the optimization of the process of publication.

d. Metadata

Currently, with the support of Sweden statistics, INSTAT has developed the MetaPlus system for maintaining structural metadata for variables and their domains for all statistical products. In terms of agricultural statistics, metadata process is in its initial stages. INSTAT is also supported by the expertise of CIM & GIZ for creating metadata of prices for agricultural products and livestock.

e. Evaluation and quality control

SIDA Project has assisted INSTAT for evaluation process and quality control and several distributed procedures are currently applied for assessing the quality of data. INSTAT has started to apply the integration of evaluation and quality control indicators in survey questionnaires. Similar questions are included in the annual survey of agricultural 2016 and calibration survey.

Anyhow, the institutions deem important the process of drafting and implementation of a strategy for the implementation of an integrated evaluation and quality control and reporting of the results of this process considering also the users opinion.

f. Institutional Capacities

In the view of both institutions (MARDWA and INSTAT), institutional capacities for agricultural system have deficiencies in terms of quantity and quality. Organizational structure of agricultural statistical system has had several changes in the last 5-years. Structural changes have been carried out to adapt to the changes of the role and responsibilities of each institution.

Up to 2012, the Ministry of Agriculture had a statistical system with 71 specialists (4 in central level and 67 established in 12 regions and 26 districts of the country). Furthermore, the government of Albania, by the Decision of the Council of Ministers Nr. 116, date 15.2.2012, has decided to centralize the Statistical system in Albania and from 2013 INSTAT is appointed to be the responsible institution for collecting, processing and publishing agricultural statistics. With this decision, MARDWA structure was reduced to 20 specialists in districts and 4 specialists at the central



level. Following this decision, INSTAT has established the Directorate of Agriculture and Environmental Statistics.

With the recent decision of the Council of Ministers No. 53, dated 21.01.2015, and the return of some statistical activities in MARDWA, the organizational structure was changed to 47 specialists distributed in 13 regional Directorates of agriculture and 3 specialists at central level. The concern that arises in this situation is: Can this structure that is relatively smaller to that prior to 2012 afford the flow of agricultural statistics activities planned and required to be implemented in the future, to satisfy the requirements of users and policymakers and in accordance with the agenda of EUROSTAT?

Beside the capacities in terms of quantity, an important concern is also the qualitative aspect of institutional capacities.

This is a common concern for all statistical system in Albania, not only agricultural. Albania does not have a university course of studies of proper genuine statisticians education

SIDA Project has had several meetings with Albanian Universities representatives to consider the opportunity and facilitate the draft of curricula of a bachelor course of studies in statistics, but it was concluded that there is not enough interest among aspiring student to motivate the start-up of such a course.

However, the SIDA project funded the establishment of a training school for the staff of INSTAT and other statistical agencies, which is expected to initiate by the end of this year. Meanwhile a training strategy is developed in order to achieve a sustainable training system for enhancing institutional capacities.

3.2 Conclusions

From consultations with stakeholders and analysis of the current situation of agricultural statistics, the summarized main issues, identified by this mission that need to be addressed in the future, are as follows:

➤ **Strengthening inter-Institutional cooperation**

This is a requirement of all institutions that are part of the statistical in Albania. INSTAT requires timely response and quality from statistical agencies, while the agencies need to have flexibility communicating with INSTAT for micro data and time series for policy making. Also, the necessity of strengthening the role of INSTAT as methodological leader for statistical agencies emerges.

➤ **Strengthening Capacities**



Institutional capacity of statistical system in general and that of agricultural statistics in particular need quantitative and qualitative strengthening

➤ ***Establishment of the farm register***

The Farm Register is a necessary and expected instrument for initiating other statistical activities for agriculture.

In establishing the farm register, special importance must be given to Unique Farm Identification Code, as a tool that can ensure the cooperation and update of farm register with existing and future registers (Livestock matriculation register, the register of agricultural supporting schemes, cadastre of vineyard etc. set up and administered by MARDWA or ARDA)

➤ ***Increase the number and frequency of agricultural surveys***

This need comes from the shortcomings identified during the process of drafting and monitoring of agricultural policies and strategies, as well as for meeting the requirements of EUROSTAT and growing demands of other users

➤ ***Strengthening administrative data***

Administrative data are a very important and accurate source for statistics. Every statistical system in every country tends to be oriented toward administrative data. In Albania, strengthening the role and use of administrative data is related with the establishment and improvement of aforementioned registers. Still, considering actual terms and what is defined as administrative data in section b.2 of this report, MARDWA needs assistance in designing a methodology for improving the quality and effectiveness of this important source of agricultural statistics in Albania. There is a need for establishment of a methodology for data collection, processing and dissemination of data that includes also instruments for forecasting and evaluation for agriculture and livestock production during and in the end of agricultural year.

➤ ***Integrated Statistical Information System***

Agriculture and other sector strategies are generally multi sectoral. For ensuring the success on their design, implementation and monitoring there is a need for a Coordination Mechanism or establishment of an Integrated Information System. This is an initiative undertaken by the current project and congratulated by stakeholders.



4. Problems encountered during the mission

No problems were encountered during the mission.

5. Recommendations for improving the sector

TO BE ANALYSIS

- ❖ After the establishment of the new statistical farm register by the agricultural census, expected after the process of calibration, a reform on agriculture statistics is necessary in terms of conducting special surveys for sensitive indicators; the sample to be objective-specific; the frequency should vary according to the objective and of course the indicators should be revised to be more compliant with EU requirements.
- ❖ Although, despite the outcome of the census calibration to move ahead statistical activities in agriculture, alternative sources must be explored and exploited such as:
 - Development of an action plan to address priority problems,
 - Improve the collection and processing of administrative data,
 - Development of metadata,
 - Analysis of the results of the annual agricultural survey conducted on the basis of population generated by the census of agriculture and calibration survey to explore the possibilities of improving and updating the database of the census and the establishment of farm register
 - Review and improve the methodologies on existing data collection
 - Design and piloting of the establishment of FADN, etc.

The focus of agricultural statistics in Albania is put on quantitative data on agricultural activity during the year. It is important to plan the production of qualitative data and statistics on socio-economic development of rural areas that can provide a deeper understanding of the impact of Agricultural and rural development policies.

- ❖ Strengthen and increase capacity in the system of agricultural statistics is a necessity. SIDA Project and the current IPA project can unite their energies to formulate a specific long-term training plan for agricultural statisticians to ensure the continuous training of experts in this field.
- ❖ Building a strategy for the publication of statistics, to address the necessary indicators to be published, their frequency, coherence, necessary explanations for users etc. is an important tool for improving the dissemination process.
- ❖ It is also necessary to draft a plan of action for the evaluation of any stage to the statistical business process and quality control.



- ❖ Improvements of statistical business processes should be guided by the framework of an Information System that needs to be drafted soon, and implemented after the establishment of all its components.

d. Plans for future missions

Future missions shall be planned as stated in the Inception report

e. Attachments (*reports, documents, presentations, etc.*)

Report on consultation document



ANNEX 4: MISSION REPORT – OGERTA ELEZAJ

Project Title Contract no.	Support to the Improvement of Statistical Information System – Albania EuropeAid/136334/IH/SER/AL
Beneficiary	Statistical Institute of Albania (INSTAT)
Contractor	The project is implemented by a consortium led by ASCENTA IT SERVICES SRL in cooperation with IN2 Zagreb
Name of Short-Term Expert	Ogerta Elezaj
Component	Component 1: Redesign and improve corporate statistical business processes; and enhance institutional capacities in the production and usage of Agricultural and Labour Force Statistics (LFS)
Sub-Activity	Sub-activity 1a2.8. Consultation with the Ministry of Social Welfare and Youth Sub-activity 1a2.9. Consultation with the Ministry of Finance
Date of mission	13-27 March 2017
Date of report	04 April, 2017

Description of mission objective

- Expertise advice and technical assistance to assess INSTAT capacities and training needs in data analyses and for data management and data integration in a diverse statistical system.
- Consultation with the stakeholder in other government institutions which provide administrative data to INSTAT to prepare Output1 a2.2: Report on consultations with emphasis on the specificity of the Labour market for Task 2.
- Assessment of Labour Force data providers capacities and the consistency between sources (part of Output 1a2.5: Report: Review of the proposed improvements and status of implementation)



List of meetings held during the mission

During the mission the following meetings with Beneficiary were held:

- Meeting with Social Statistics directorate

Participant	Email	Role
Ledia Thomo	lthomo@instat.gov.al	Director of Social Statistics
Pranvera Elezi	pelezi@instat.gov.al	Head of Labour Market and Salary Statistics Unit

- Meeting with Ministry of Social Welfare and Youth representative

Participant	Email	Role
Genta Qosja	gentaqosja@mpcs.gov.al	Director of Employment Policy Unit

- Meeting with National Employment Service representative

Participant	Email	Role
Ms. Etleva Gjelas	etlevagjelas@yahoo.com	Statistics Sector

8. Results achieved during the mission

- Report on consultation with stakeholder.
- General overview of survey process in INSTAT with a special focus on LFS survey
- General overview of data analyses and data dissemination in INSTAT with a special focus on LFS survey
- General overview of data sources, formats and software used during data processing in INSTAT

Report on Consultation with stakeholders

Consultation with the stakeholders: Ministry of Social Welfare and Youth, National Employment Service, and SIDA project

Consultation with LFS team in INSTAT and other relevant data providers about training needs in the fields of LFS

LFS TIME SERIES ANALYSES



INSTAT introduced a Labour Force survey in 2007 and this survey was carried out on an annual basis till 2010. In the survey are selected individuals aged 15 years and over. Working age population is the population between 15-64 years old. All individuals aged 15 years and over in the selected household are subject of labour force survey. Based on EU regulations and 5 years official statistical program, INSTAT introduced for the first time a quarterly survey in April 2011 with a rotating sample design. The same selected household is interviewed for 5 consecutive quarters. Also in this process there is a replacement of one-fifth of household in every quarter. On the other hand, administrative data is collected and provided to INSTAT based on the 5 years Official Statistical Program. The employment data of non-agriculture private domain are collected and provide to INSTAT by National Employment Service. The main source of these data is the regional employment offices and the General Directorate of Taxation. Two sets of data are used: data on non-agriculture employment in the private sector compiled by the National Employment Service (NES) based on tax returns of establishments; and data on jobseekers registered at the NES.

One important component in the labour force analyses is the agriculture sector employment for which there is not any administrative data source available. These data are estimated based on the information collected by the LFS survey.

Regarding the unemployment figures, the information is provided to INSTAT by Ministry of Social Welfare and Youth. As the data sources and the information available is increased further progress should be done in strengthen the INSTAT staff in making in depth analyses of labour market data. INSTAT Labour Market and Salary Statistics Unit gave detailed information about the survey process and also, about tabulation programme used for the compilation and publication of results. LFS unit contains wealth information that should be tabulated using time series comparison, gross flows measurements and increase the number of indicators produced. INSTAT staff emphasise the need for calculation of labour cost indexes on a quarterly basis and econometric modelling.

The representatives from Ministry and National Employment Service require actively participating in INSTAT training for increasing their capacity in data analyses and in using SPSS software for data processing.

My recommendation is to increase the LFS staff capacities via a time series training which should effectively address the growing and diverse needs for skills and competencies in INSTAT in both its content and methods for the following topics:

- Characteristics of time series data
- Stochastic models
- Forecasting of several time series



- Examples and applications to the LFS survey (SPSS examples)

INSTAT Data Management System

INSTAT main data sources are surveys and administrative sources. INSTAT continues to collect mostly of the data in a traditional way via face to face interviewing using paper questionnaires or laptops as medium to collect the data (CAPI). Different modern ways of data collection such as CATI or web based are testing but not fully implemented. The data of different surveys are saved in different formats. Mostly of the social surveys data such as LFS, HBS, Living Standards Measurement Survey (LSMS), European Union Statistics on Income and Living Conditions (EU SILC) are processed using SPSS software, censuses data using SAS/SPSS and the economic surveys using MICROSOFT ACCESS/SPSS and some price indexes using SQL Server. As the data are coming from different sources and are in different formats it is very important for INSTAT staff to get trained in data processing and data integration.

The processes such as editing and imputation, logical/physical integration of administrative and statistical data, statistical matching and record linkage are crucial processes for which INSTAT staff should be trained to get further theoretical and practical background.

The process of record linkage is seen by the sector as crucial process used to match record coming from different sources and for longitudinal linkage of individual records from different LFS rounds. Also, this record matching is very important because the quarterly data transmitted to EUROSTAT should be sending with longitudinal key as it will allow Eurostat to produce flow estimates for the country.

But, with the new trends in official statistics, there is a great potential of new data sources which resides in real time generated at a lower cost compared to the official data sources. Big data, characterized as data sets of increasing volume, velocity and variety, is unstructured data and harvesting the information from the big data and incorporating it into a statistical production system is not easy. It needs to adequately address issues to pertaining to methodology, representativeness, quality, technology, data access, legislation, privacy, management and finance, and provide adequate cost-benefit analyses.

As many European statistical offices are moving towards the use of Big Data profiting from the advantages of using them in their statistical system, INSTAT also has to increase capacities in this field.

My recommendation is to increase the staff capacities via a data management training which should effectively address the growing and diverse needs for skills and competencies in INSTAT in both its content and methods for the following topics :

- Data processing and data integration
- Quality and access to data
- Metadata System
- Examples of Big Data uses in European Statistical System

Also, as a request from other institutions, we recommend including in the training sessions participants from different INSTAT units dealing with statistics production (not only LFS unit representatives) and participants from other institutions working in statistical units.

Regarding the software used during LFS data processing, there is a need to build statistical capacity for the staff to develop SPSS parameterized procedures and increase skills in writing their own macros to increase productivity and improve LFS data management. The methodology unit is using SAS procedures during weighting phase and also some procedures available by EUROSTAT for checking micro data before transmitted are written in SAS, training in SAS are important for the sector. The staff has limited knowledge in SAS. Also, further discussion are needed to be done with the IT team for the software licenses plan, because SAS/SPSS are not open source and in case of limited budget training in R (open source statistical package) should be considered for the staff. At the current moment, it is not decided yet if SIDA project will provide INSTAT with the required licenses for SAS/SPSS software.

The staff of INSTAT, particularly, Ms. Ledia Thomo and Ms. Pranvera Elezi were extremely cooperative and provided all the necessary information as well as their valuable insights.

9. Problems encountered during the mission

No problems have been encountered during the mission.

10. Recommendations for improving the sector

The general recommendation for INSTAT staff:

- General capacity building regarding the theoretical and practical knowledge in time series analyses with a focus on stochastic models and seasonal adjustments.
- More extensive training in SPSS/SAS (parameterized macros) using real LFS data sets for data management and time series analyses;
- General capacity building regarding the Data Management of Statistical



Information System including data sources linkage, data processing and data quality.

- Active participation in all the training sessions of representatives of National Employment Service and Ministry of Social Welfare that are in charge of labour statistics.

Plans for future missions

Future missions will be planned as stated in Inception report

11. Attachments (*reports, documents, presentations, etc.*)

- Report on consultations with the stakeholders
- The list and content of proposed comprehensive training programme for LFS statisticians (trainings and workshops)